

Panoramica sulla compressione video: formati e caratteristiche (2)

Dopo aver visto, con una certa dovizia di particolari, caratteristiche e peculiarità di alcuni formati di compressione video più significativi, continuiamo il discorso con altri esempi, anch'essi altrettanto significativi ed importanti. In questa seconda parte ci occuperemo di Indeo (Video 1 e RLE), CinePak, QuickTime e Fractal Transform

di Massimo Novelli

Intel Indeo (Intel Video Technology)

Parlare di Indeo senza occuparci di Microsoft Video For Windows è impresa ardua; l'ambiente integrato audio/video creato dalla casa americana ormai 2 anni fa, e recentemente aggiornato con modifiche, ci offre un ammirevole «ombrello» multimediale sotto cui molte delle attuali implementazioni, provenienti soprattutto da case diverse, prendono posto, evitando altresì di dover scrivere software per lo scopo.

Considerato dalla stessa Microsoft alla stregua di una vera e propria estensione di Windows 3.1 (e che vorremmo vedere farne parte integrante nella prossima release del sistema operativo), Video For Windows, con i suoi ambiti di cattura e di editing offre una comodissima soluzione, a vari livelli e con potenzialità ben calibrate. A cominciare, e non poteva essere altrimenti, con i formati di compressione consentiti, che vanno dal già citato Indeo (ora nella versione 3.0), al CinePak della Super-Mac, al Video 1, sempre Microsoft su licenza Media Vision, all'RLE.

Le caratteristiche principali di esso, al di là della gestione dei file con il software in dotazione, è nella capacità di generare file audio/video, e quindi nel gestire il codec, in modalità cosiddetta «scalare», cioè nell'adeguarsi alla potenza della macchina ospite, sia nella cattura come nel playback, dei famosi file .AVI (Audio Video Interlace). Per fare un esempio, ciò che sarà solo dignitoso su una macchina classe 386 sarà invece molto soddisfacente su un 486 e così via, sia in termini di quadri/sec., e quindi di flusso, sia in termini di dimensioni finestra (a cui, ovviamente, contribuisce anche la velocità intrinseca della scheda video VGA in dotazione).

Parlare precipuamente di Indeo però, almeno nella versione solo software come quella presente in Video For Windows, non sarebbe corretto senza fare riferimento a quella assistita da soluzioni hardware associate ad esso, come quella ActionBoard II (scheda di cattura/playback Intel/IBM) di cui spesso si parla. Poiché la «vera», almeno in efficienza, implementazione di Indeo (e soprattutto del formato DVI) passa proprio per questo tipo di hardware, mediante l'impiego di chip appositi (il famoso i750B) oppure nell'altrettanto famosa Intel SmartVideo Recorder, altra produzione audio/video già vista su queste pagine, che adotta lo stesso hardware sebbene dedicato in modo leggermente diverso dalla precedente.

In dettaglio, quindi, la differenza tra le due si estrinseca nella completa gestione del chip i750 in un caso, mediante le possibilità di essere micro-pro-

grammato (sia in modo proprietario, e diverse case lo fanno tuttora, sia in modo standard) e nel secondo caso, apparentemente più modesto, curiosamente avremo la parte hardware-playback misteriosamente «limitata» e non influente allo scopo, ragion per cui si ricorre, al solito, ai codici software.

In generale, comunque, il playback dei file Indeo (con codec software) sarà di almeno 15 fotog./sec. in un quarto di schermo, con colore a 24 bit, su PC 486 a 33 MHz o più.

Le varianti operate in VFW 1.1, rispetto al precedente, offrono ulteriori miglioramenti nel codec Indeo, a sua volta passato dalla 2.1 alla release 3.0. Una maggiore velocità di intervento, unita ad un più efficiente controllo della compressione, sia in rapidità che accuratezza, sono le prerogative principali.

È ovvio, stiamo sempre parlando di soluzioni «only-software», con tutte le conseguenze del caso, ma la differenza



Figura 1 - Le opzioni di compressione di Indeo 2.1, con l'about aperto; superato ormai dalla versione 3.0 e 3.1, Indeo continua a dare ottimi risultati.

si nota comunque. Fondamentalmente di tipo «lossy» e a codifica simmetrica/asimmetrica, Indeo ha delle caratteristiche, in termini di algoritmi di compressione, molto simili alle tematiche affrontate dallo standard MPEG, seppure con diverse differenze, soprattutto nell'utilizzo dei cosiddetti «key frames» (fotogrammi chiave); in sintesi, detto key-frame sarà un fotogramma, scelto per un dato intervallo, del quale verranno conservate tutte le informazioni ed in base a simili dati, di tutti i frame posti tra due key-frame, verranno registrate solo le informazioni relative ai pixel che hanno variato colore. Più sarà alto il suo valore meno dati verranno registrati e minore risulterà l'occupazione di memoria, altrimenti viceversa le variabili si invertiranno.

Il suo anello debole, se così si può chiamare, sarà nella calibrazione ottimale dei parametri di intervento, sulla cattura video principalmente (e quindi sul file originario), oltre che nel playback conseguente, dopo aver attuato la compressione. Esasperare qualche parametro a scapito di altri, può risultare deleterio se non congruente al massimo; si potrà comunque operare dai 15 ai 30 fotog./sec., generando un flusso di circa 1.2-4.8 Mbit/sec. e con finestre di 160 x 120 oppure 320 x 240 pixel. I prodotti Indeo hanno dalla loro una qualità finale decisamente buona, anche se perfettamente migliorabile a fronte di codifiche software, crediamo, su cui ulteriormente lavorare.

Microsoft Video 1 e RLE

Gli altri codec facenti parte di Video For Windows, il cosiddetto Video 1 e il formato Run-Lenght-Encoder (RLE), si possono catalogare a pieno merito degni antagonisti del precedente, almeno il primo (per il secondo, il suo utilizzo è limitato).

Il primo (su licenza Media Vision), è abbastanza simile all'Indeo e offre prestazioni dignitose e rispettabili, a fronte di una serie di scelte qualitative che lo paragonino almeno all'Indeo. La sua efficienza è comunque buona, anche se genera file leggermente più ingombranti del suo concorrente. Un buon compromesso tra qualità finale, data rate e frame/sec.

L'RLE, il più semplice attualmente

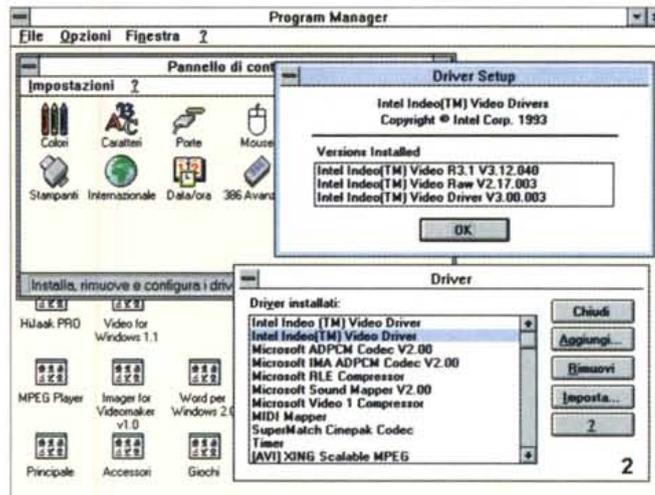


Figura 2 - Siamo in ambito Pannello di controllo di Windows, e appena dopo aver installato Video for Windows 1.1; in evidenza i driver video presenti nella confezione.

dei codec in VFW, opererà in modo molto simile, per fare un esempio, alla trasmissione fax, rimpiazzando sequenze di pixel identici con singoli «treni» di codici e con conteggi ripetuti. Generalmente, infatti, esso lavora bene su file di animazione o su ogni cosa costruita

o disegnata, in sintesi generata, da computer. Questo tipo di immagini, di solito, contengono aree più o meno grandi di colore «solido», a tinta unita, e ne verrà permessa un'adeguata compressione. Ma se si proverà a processare un video, con l'RLE, esso mo-

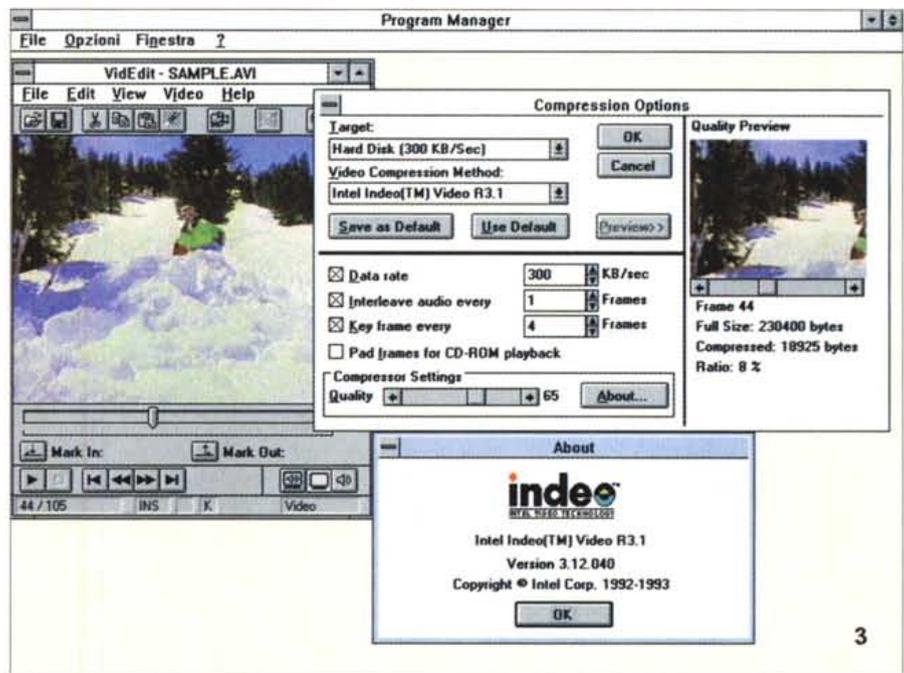


Figura 3 - La nuova versione di Indeo, come abbiamo detto la 3.1, offre ulteriori possibilità in compressione, dotata com'è di una più spiccata «scalabilità».

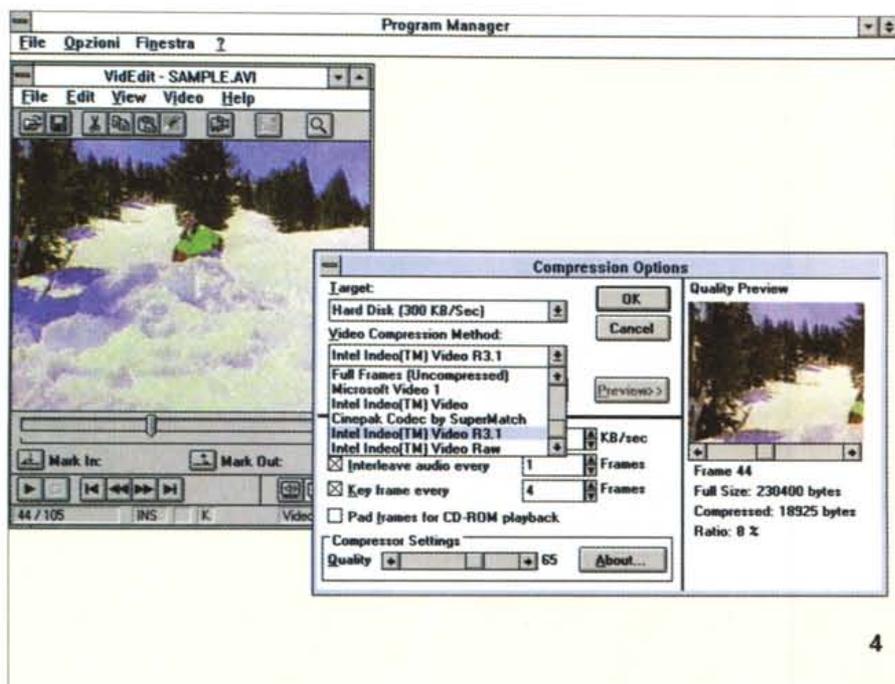


Figura 4 - La nuova serie di codec presenti in Video for Windows 1.1; c'è ne per tutti i gusti, ma ne manca ancora qualcuno, uno per tutti quell'MPEG in cui tanto speravamo.

strerà subito la corda, con spesso una generazione file compresso più grande dell'originale, per non parlare del degradamento della qualità.

Compressione, quella dell'RLE, tipicamente adatta ad animazioni grafiche e a slide show di grafica generata da PC, non di «live video».

Tornando a parlare di Video For Windows, invece, nella sua ultima versione avremo le possibilità di operare in almeno due diversi modi aggiunti, sia in modo «Full Frame» (non compresso, generando un salvataggio dei singoli frame come .DIB), che in Indeo Raw (ad uso Intel Smart Video Recorder, in formato YUV non compresso).

Ma diverse implementazioni di altre case, come abbiamo detto all'inizio, prendono posto nella struttura base di VFW; a cominciare dalle librerie presenti in Media Player, il modulo playback di VFW (e a cui faranno riferimento gli specifici driver MCI installati).

Una per tutte, e già vista su queste pagine, l'implementazione di decoder software MPEG della Xing Technology, che consente di andare in play, appunto dal Media Player, con file in standard MPEG, previa codifica in questo formato proveniente da hardware adatto allo scopo. Per usare una tale alternativa, poi, sarebbe curioso chiedere a Microsoft il perché non abbia realizzato, nell'ultima release di VFW, possibilità

anche in tal senso, vista ormai l'ampia accettazione dello standard, in diversi ambiti e con egregie soluzioni.

SuperMac CinePak

E veniamo senza indugio ad un altro codec, molto quotato, quel CinePak della SuperMac, già adottato praticamente da tutta l'industria del multimediale, da Apple a Microsoft, 3DO, Atari, Sega, Creative Labs e Cirrus Logic, solo per fare qualche nome.

La tecnologia del codec di CinePak era stata inizialmente sviluppata per il QuickTime di Apple ed è diventata oggi quasi uno standard per l'intera industria, vuoi perché molte case hanno già acquisito licenza d'uso, vuoi perché oggettivamente offre risultati egregi quasi senza sforzo.

Tipicamente, la sua finestra video utilizzerà i classici 320 x 240 pixel e potrà mostrare da 12 a 30 fotog./sec., secondo due variabili fondamentali; la velocità della CPU a cui opera il PC e la velocità dell'eventuale CD-ROM collegato al sistema (per un playback di materiale registrato in simile formato). In ogni modo, nel peggiore dei casi, esso sarà comunque in grado di mostrare almeno 12 fotogrammi al secondo, mentre i dati verranno compressi, mediamente, con un rapporto di circa 20:1; abbiamo detto mediamente poiché il reale rap-

porto di compressione dipende dalle caratteristiche delle immagini, in particolare dalla quantità di pixel ripetuti identicamente nella sequenza video. I dati considerati inutili verranno impietosamente scartati, con risultati comunque di tipo «lossy», rendendo così lo spazio necessario per contenere un minuto di video variabile da 3 a circa 10 MByte di memoria.

CinePak è basato su soluzioni cosiddette di quantizzazione vettoriale, con risultati particolarmente buoni per dati a 24 bit, e necessiterà, per la compressione, di macchine particolarmente veloci, potendo comprimere ciascun fotogramma in circa 10-20 secondi; e traducendo tutto ciò, per esempio, in un'ora di compressione per un minuto di video a 15 fotog./sec. È infatti estremamente «lento» (almeno nella versione solo-software), ragion per cui diverse case stanno già considerando l'idea di implementarne una soluzione hardware che offra tempi migliori.

Crediamo sia, attualmente, una buona soluzione da verificare con tranquillità, anche se l'uso è in qualche caso estenuante.

Apple QuickTime

Che cosa si potrebbe dire di QuickTime Apple? Giunto alla versione 2.0 (di cui sono state date anticipazioni qualche mese fa), si potrebbe dire che è l'estensione multimediale principe del S.O. Mac. Esso infatti integra tipi di dati, come audio, video e animazioni, in modo dinamico, su base temporale, nelle applicazioni dedicate, considerandole come «movies». Infatti, come questo termine implica, QuickTime sarà in grado di gestire eventi che avvengono nel tempo in un modo molto simile a quello che QuickDraw può fare con le immagini fisse. Ad ulteriore riprova, comunque, QuickTime offrirà un approccio standard al modo di comunicare con le applicazioni, tipica filosofia Apple, per cui non si dovrà essere necessariamente un professionista per gestire simili procedure; sarà «hardware indipendente» e ne è stata già immessa sul mercato una versione per Windows.

La sua architettura consiste in quattro componenti principali; il «system software», i «files format», compressori/decompressori (codec) e l'interfaccia standard di comunicazione. Il system software ha tre sub-componenti: il Movie Toolbox, per la creazione, l'editing ed il playback dei movie; l'«Image Compression Manager» o ICM ed il «Component Manager».

Un cosiddetto «movie» sarà considerato un contenitore di tutti i dati appena

menzionati, dinamici, e potrà essere usato per presentazioni slide show, animazioni, montaggi di immagini e suono, ecc. Esso conterrà riferimenti a gruppi di dati omogenei tra loro, come appunto il video e l'audio, che saranno organizzati come «Tracks»; ma esse non conterranno i veri e propri dati, ma solo quei riferimenti necessari a ritrovare i file immagazzinati sui «media memory» (HD, dischi ottici, ecc). QuickTime dovrà solo sincronizzare queste tracce quando andrà in playback di un evento.

D'altro canto, Apple ha esteso di molto il formato file PICT, che ora consente di essere compresso e mostrato secondo schemi ben definiti; l'utente infatti potrà comprimere un'immagine usando uno qualsiasi degli standard di compressione forniti nel Component Manager e decomprimerne il risultante, una volta installata l'estensione QuickTime.

Il tempo è un concetto fondamentale nei movie di QuickTime; ognuno di essi infatti ha una scala dei tempi, una base dei tempi ed un sistema di coordinate temporali. La scala dei tempi definisce le unità di misura e la durata dei movie, la base dei tempi è il valore corrente del

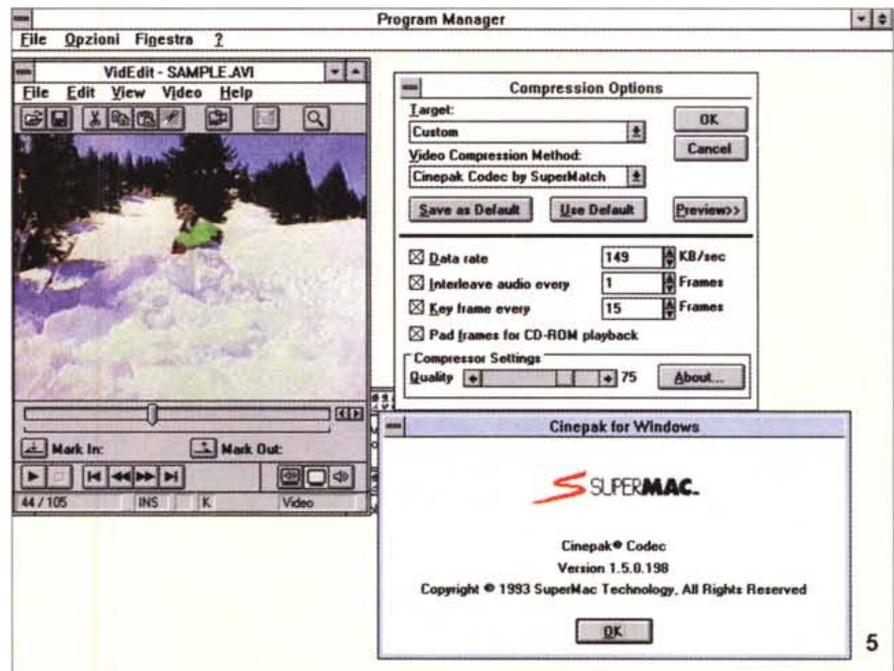


Figura 5 - Stiamo settando la codifica CinePak in Video for Windows; al di là della sua lentezza di intervento, è uno dei codec più riusciti e egregi, potendo generare file molto compatti e poco «lossy».

Metodi di compressione video

Metodo	Frame rate frames/sec.	Data rate Mb/sec.	Risoluzione pixel	Audio Sincronizzato	Hardware speciale	Compressione	Qualità
Ultimotion (IBM OS/2)	15	1.2	160 x 120	si	no	simmetrica	3
CinePak (QuickTime 1.5)	15-24	1.2-4	320 x 240	si	no	asimmetrica (150:1)	3
Indeo (VFW)	15-30	1.2-4.8	160 x 120 320 x 240	si	no	simmetrica o asimmetrica	3
PLV (DVI)	30	1.2	256 x 240 (640 x 480)	si	i750	asimmetrica (60:1)	4
MPEG-1	30	1-4.7	320 x 240 (640 x 480)	si	C-Cube CL 450/950 CLM 4500	asimmetrica (15:1)	4
MPEG-2	30	2-20	720 x 480	si	C-Cube CLM 4600	asimmetrica	5
Motion JPEG	30	4.8-10	160 x 120 640 x 480	no	C-Cube CL 550/560	simmetrica	3-4
TrueMotion (Horizon Tech.)	30	4.8	384 x 480 (768 x 480)	si	i750	asimmetrica (10:1)	4-5

Le risoluzioni in pixel sono valori tipici o massimi, come d'altronde i frame rate ed i data rate. Le risoluzioni in pixel, indicate in parentesi, sono immagini full-screen ottenute mediante interpolazioni orizzontali e raddoppio di linee. Le caratteristiche dei pixel, nel PLV, hanno un rapporto di aspetto di 5:4 (e non il classico 4:3 televisivo). La qualità finale è in scala da 1 a 5, cioè verso la migliore.

I dettagli dei più comuni schemi di compressione video in uso attualmente; i livelli di qualità sono relativi e dipendenti dai rapporti di compressione usati. Lo schema TrueMotion è di recentissima entrata sul mercato.



Figura 6 - Un altro pre-tendente, quel Video 1 Media Vision/Microsoft, offre delle buone caratteristiche di fondo, adatto per tutti gli usi.

In estrema sintesi, e non potendo andare più in dettaglio sulle capacità di QuickTime perché troppo estese per questo ambito, la sua potenza, e soprattutto le sue estensioni ad architettura aperta, ne fanno un prodotto estremamente flessibile e oltretutto facile da usare, merito anche della filosofia Apple che privilegia quanto più possibile l'utente; la versione 2.0, che aggiunge componenti come l'MPEG, consente ulteriori vantaggi nella gestione e nelle capacità multimediali della linea Apple Mac. Un prodotto veramente potente e prepotente (rispetto alla concorrenza).

Iterated System «Fractal Transform»

La compressione di immagini, per via frattalica, non è decisamente nuova né così sperimentale come si potrebbe supporre. Al momento dedicata alle sole immagini fisse (data l'estrema complessità di processione), e non ancora a punto per il video, nei fatti offre il massimo della qualità finale, con ingombri ridicoli dei file generati, partendo da originali esageratamente estesi è stata messa a punto dal matematico americano Michael F. Barnsley, fondatore tra l'altro della Iterated System detentrica della tecnologia.

In breve sintesi, parlando di frattali, di essi si potrà dire che si tratta di «im-

tempo lungo un vettore che definirà la direzione (in avanti o indietro) e la velocità dello stesso, che verrà estratta dal sistema di coordinate temporali, le quali essenzialmente non sono altro che il vettore time base, su cui «scorreranno» i dati.

L'ICM di QuickTime permetterà alle applicazioni, con compressione e decompressione di immagini, di essere «device and algorithm independent», a causa della sua architettura aperta, che consentirà sia di supportare diversi tipi di algoritmi, sia di interpretare ed eseguire schemi di compressione senza fatica. In questo modo, l'ICM sarà in grado di permettere a tutto il software compatibile QuickTime di usare le comuni interfacce utente per operazioni di compressione/decompressione, su immagini o sequenze di immagini, e di permettere alle stesse applicazioni di usare dati provenienti da altre.

Gli algoritmi di compressione presenti in QuickTime (almeno nelle versioni 1.5/1.6) sono di varia natura e scopo; i tre principali sono il JPEG (photo compressor) che QuickTime usa per le immagini PICT, il video compressor, proprietario Apple, che permette dei veloci tempi di decompressione pur mantenendo una qualità finale decisamente buona, e l'animation compressor, dedicato per animazioni e grafica generata da computer.

Parlare del primo ci sembra superfluo, mentre del secondo si potrà dire che permette un display interattivo di immagini a 24 bit, è in grado di processare immagini video ad 8, 16 e 24 bit, con supporto di entrambe le compressioni in modo spaziale e temporale; il terzo componente è basato invece su algoritmi ottimizzati per animazioni in

genere e per videografica generata da computer.

In dettaglio, QuickTime 1.5 potrà mostrare a video movie di 160 x 120 pixel, 8 bit colore, da 24 a 30 frame/sec., oppure movie 320 x 240, da 12 a 15 frame/sec. su un Mac LC II; un più veloce Quadra 950 ci consentirà invece di vedere movie di 640 x 480, 8 bit colore, con suono mono ad almeno 24 frame/sec., appena un passo sotto al «full motion video».

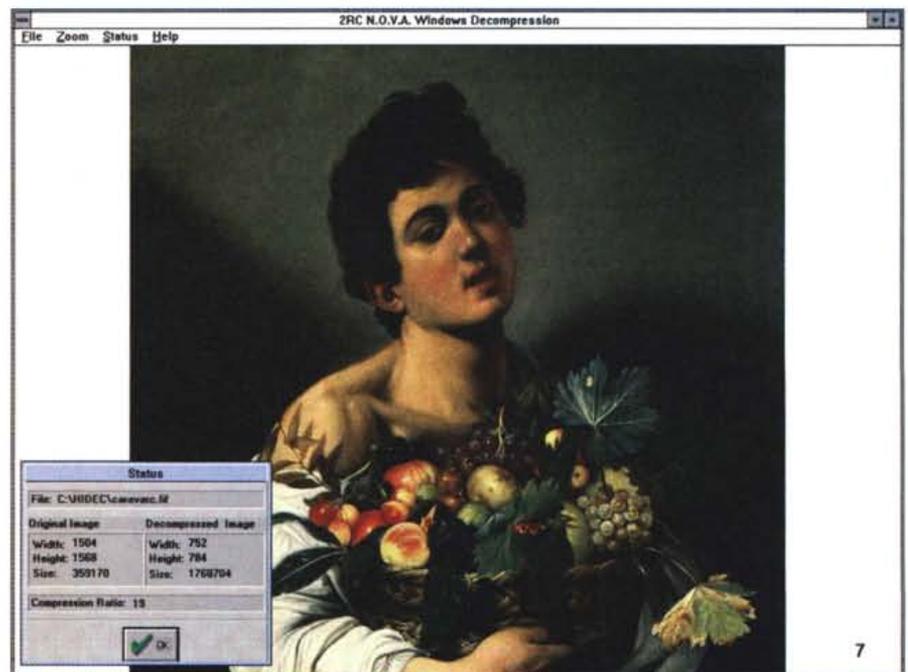


Figura 7 - L'emergente Fractal Transform, per ora dedicato solo allo «still-image», è dotato di una potenza di compressione mai vista, unita a caratteristiche di ottima qualità, scalabilità e performance.

magini infinitamente ingrandibili che possono essere prodotte da un relativo piccolo set di istruzioni e dati». In una immagine del genere, più si andrà in «zoom» più si avranno dettagli della stessa, a differenza del classico bit-map in cui, dopo qualche ordine di grandezza, noteremo con fastidio la «pixelation» dei blocchi facenti parte l'immagine.

In dettaglio, si potrà parlare del concetto di «trasformazione affine», centrale nella compressione di immagini per via frattalica. Essa infatti è una funzione matematica comprendente gradi e combinazioni di rotazioni, scalature, distorsioni e traslazioni dei pixel, in uno spazio dimensionale -n. Si produrranno quindi delle matrici con cui verrà analizzata l'immagine, dandone la rappresentazione matematica più consona alla sua natura.

Diversi teoremi ci vengono in aiuto nel capire come espressioni matematiche possano rappresentare un'immagine «reale», tra gli altri il «Collage Theorem» ed il «Chaos Game», ragion per cui Barnsley ebbe a suo tempo l'intuizione di considerare che tutte le immagini del mondo reale sono ricche di ridondanza affine, e che questo significa quantomeno che, mediante trasformazioni affini, saremo in grado di rappresentare, anche con meno dati a disposizione, la stessa immagine senza apparente perdita di dettagli.

Il primo passo nella compressione sarà di dividere l'immagine in regioni, non sovrapposte tra loro, e per ogni regione considerata la tecnica dovrà scegliere le trasformazioni affini, in 3D, che più da vicino rappresentano i dati originari. Esse non solo deformeranno e varieranno la parte di immagine, ma interverranno anche sul contrasto e sulla luminosità della stessa. Ogni trasformazione sarà così descritta da propri coefficienti.

Un file FIF (Fractal Image Format) sarà così scritto, e conterrà un «header» con informazioni sulle specifiche scelte operate nelle regioni, seguita da una lista di coefficienti affini considerati per ogni regione. Tale processo genererà file che sono indipendenti dalla risoluzione dell'immagine originale; per dirla in termini molto «terra terra» si sarà così trovata «un'equazione» dell'immagine.

Ma tutto ciò implica tempi di codifica a volte molto lunghi, ragion per cui nelle applicazioni commerciali della Iterated System, sono stati introdotti dei limiti pratici sul «range» delle possibili regioni considerate e nelle trasformazioni permesse. In uno di essi, il «Poem ColorBox», per esempio, il

Figura 8 - Uno schema riassuntivo delle prestazioni di alcuni dei codec considerati, con in evidenza la qualità video rispetto al «data rate» e alle richieste di memoria di massa.

